DOCUMENT RESUME

ED 206 739 TM 810 694

AUTHOR Hyman, Cynthia; Tobias, Sigmund
TITLE Feedback and Prior Achievement.
PUB DATE Oct 81
NOTE 47p.; Paper presented at the Annual Meeting of the Northeastern Educational Research Association (Ellenville, NY, October, 1981).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Academic Achievement; *Feedback; Intermediate Grades; Judaism; *Learning Processes; Literature Reviews; *Pretests Posttests; *Prior Learning; *Programed Instruction; Reinforcement

ABSTRACT

The hypothesis that feedback in programmed instruction is an important variable in the learning of novel, but not familiar, content was investigated. A linear, constructed response program dealing with the Sabbath rituals in the synagogue was chosen due to wide variability in student familiarity with this topic. Subjects were randomly assigned to one of three treatment groups. Feedback was experimentally manipulated by presenting knowledge of results simultaneously with each frame, only after the response had been made, or not at all. Pretest scores measured prior familiarity, and posttest scores were used as the dependent achievement measure. Regression analysis indicated a number of main effects on achievement for prior achievement and for treatment. The No Feedback group scored slightly higher than the Feedback group, and the Simultaneous Feedback group scored lowest. The predicted interaction between treatment and prior achievement was not supported. (Author/GK)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED206739

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Feedback and Prior Achievement

Cynthia Hyman
Graduate Center

Sigmund Tobias
City College

City University of New York

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

S. Tobias

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 810 694

Paper presented at the annual meeting of the Northeastern Educational Research Association, Ellenville, New York, October 1981.

ABSTRACT

This study investigated the hypothesis that feedback in programed instruction is an important variable in the learning of novel, but not familiar content. A linear, constructed response program dealing with the Sabbath rituals in the synagogue was chosen due to wide variability in student familiarity with this topic. A total of 132 subjects, 84 from a Jewish Day Camp, and 48 from a Hebrew Afternoon School, were randomly assigned to one of three treatment groups. Feedback was experimentally manipulated by presenting knowledge of results simultaneously with each frame, only after the response had been made, or not at all. Pretest scores measured prior familiarity, and posttest scores were used as the dependent achievement measure.

Regression analysis indicated a number of main effects on achievement for prior achievement and for treatment. The "No Feedback" group scored slightly higher than the "Feedback" group, and the simultaneous feedback group scored lowest. The predicted interaction between treatment and prior achievement was not supported, perhaps due to the motivational problems in a large part of the sample.

FEEDBACK AS A REINFORCER

One of the principle advantages attributed to programed instruction has been its ability to provide frequent reinforcement to learning via immediate feedback. Skinner advocated the use of mechanical devices in order to "reinforce the student for every correct response" (1968,p.39), and to insure that "reinforcement for the right answer is immediate"(1954, p. 95). This assumption that feedback in a learning program serves as a reinforcer has been widely adopted, and feedback within currently available instructional programs ranges from that of providing some indication as the correctness or incorrectness of a response, to programs where answers are provided, and sometimes even explained.

Research studies, however, have not confirmed that feedback serves as a reinforcer,or that it increases learning. In a review of the literature, Anderson (1970) notes that:

> To the behaviorist, the most shocking result of programed instruction research is the finding that programs teach as much, or more when immediate 'reinforcement' is omitted.(p. 356)

A survey of research on the effects of feedback on learning indicates that results have been equivocal with most studies showing no significant gains due to feedback availability.

Non-significant Feedback Effects

Krumboltz and Weisman (1962) studied the effects of varying schedules of feedback, basing their research on the assumption that "in programed instruction knowledge of results may be considered a form of reinforcement" (p. 250). Using programed workbooks on educational measurement, treatment groups were randomly assigned to receive feedback continuously, on a fixed ratio, or variable ratio schedule, or no information as to the correct answer.

Results did not confirm their main hypothesis. Scores on a completion type criterion posttest administered immediately after subjects finished their programed workbooks, showed no differences in learning due to the variation of feedback availability or schedule. The only difference between the groups was in error rate on the programs; the higher the ratio of feedback the fewer program errors, but there was no concomitant increase in learning.

Feldhusen and Birt (1962), utilizing a program dealing with teaching machines varied feedback availability and the type of teaching device used, simultaneously within one experimental arrangement. The program was varied by presenting feedback immediately, delayed, or not at all, or by presenting the same program in paragraph form. The manner of presentation was varied by the utilization of either a mechanical teaching device, or a manila

folder device which contained a feedback flap which could be opened and closed by the student. There were a total of nine groups of college students who were randomly assigned to one of nine variations in design.

Learning achievement, as measured by post-test scores, was high within all groups receiving the program, regardless of variations. There were no significant differences attributable to either variations in feedback presentation or teaching device used despite the fact that the manila folder device utilized allowed for the "simultaneous view of the stimulus frame and the correct responses" (p. 463).

Feldhusen and Birt (1962) conclude that the favorable learning effects may have resulted from "the care in preparation of the material as a program" and "from the constant attention to behavioral objectives" (p. 463). Based on their results, they question the need for "frame-by-frame feedback as offered in the standard linear program" (p. 466). Further research was advocated to test the effectiveness of these same variations in program presentation when learning sessions covering several days were necessary, as well as to test the effectiveness of other variations in program presentation.

Goldstein and Gotkin (1962) reviewed eight studies reported in the literature which compared machine and programed textbook presentation of self-instructional material. Their orientation was a practical one, "to assess the value of machines and text-books for program presentations" based on the "mastery of the program subject matter" (p. 34-35). They noted that the essential difference between the machine and workbooks was that the machine constrained the learners so that they could not see an answer beforehand nor alter the response after the correct answer had been revealed. In essence these studies investigated the effectiveness of feedback following rather than preceeding or occurring simultaneously with the response.

The eight studies reviewed varied with respect to topic (English, math, science, logic, psychology), age of the subjects (10 years to adult), length of the program (one to twenty sessions) and response mode (constructed and multiple choice). Results, however, were markedly uniform, with no significant achievement or attitudinal differences between machine and programed text presentation of the same material. The only reported difference between the experimental groups was a time saving in favor of the programed text group. In four of the five studies in which time to complete the

program was a variable, statistically significant differences were noted between the presentation modes, with time savings for programed text presentation varying from approximately 10 to 40 percent. Noting this time saving Goldstein & Gotkin (1962) recommended research to evaluate "the 'cheating' factor inherent in programed texts" (p. 35), to determine whether this serves as an asset or hindrence to learning.

Hough and Revsin (1963) compared the effects of providing feedback as well as the possible advantages to be accrued by using a teaching machine as opposed to programed texts. Ninety college students who were prospective teachers were given a linear constructed response program on "The Contemporary Secondary School" as part of a course requirement.

Results indicate no differences between the two programed text groups, one receiving immediate feedback of the correct answer and the other receiving no feedback. There were also no differences between the groups working on the machine and the group using the programed texts, when both received immediate feedback. The authors do note that their results may have been affected by the extremely low error rate in the program, which only averaged two percent. Nevertheless, the notion that

providing knowledge of the correct response in programed lessons is a definite advantage was, again, challenged.

Moore and Smith (1961) studied the effectiveness of immediate feedback in an attempt to deal "with the assumption that information on the correctness of each response must be provided to the student with a minimal delay"(717). The experiment was conducted in regular sixth grade classes over a 14 week period, with students within classes randomly divided into experimental and control groups. Programed instruction of their spelling words were presented to both groups for 15 minutes daily. The experimental group was given an answer booklet containing the correct response, and instructions to "check the accuracy of each answer immediately after it had been written" (p. 718). The control group received the identical program without the answer books.

Results on post tests conducted two days later showed no overall differences between the two groups. Unexpectedly, however, the control groups were consistently higher in spelling achievement and on one unit (that used during the 14th week) this difference reached significance at the .05 level.

Moore and Smith (1961) hypothesized that "since programmed units were used in a text form, cheating may have been a problem" (p. 719). They therefore conducted a second study using machines in order to provide more adequate controls. Results were replicated with the control group again scoring slightly higher in spelling achievement, but in no instance was the difference in scores significant at the .01 level (.05 level was not reported).

Since the results from neither experiment indicated any gains from providing feedback, the authors questioned the need for providing knowledge of results. Instead they hypothesize that feedback may sometimes result in over-prompting, and that "over-cueing may reduce the active participation of the student, thereby reducing learning efficiency" (p. 722). They maintained that their experiments supported claims for use of programed materials in class, since achievement level in both groups was high, without the assistance of the classroom teacher.

In a follow-up study, Moore and Smith (1964), compared the effects of varying types of knowledge of results, machine and textbook presentation, and multiple choice and constructed response, all within one experiment. Subjects were 220 college students randomly assigned to experimental groups during 15 weeks of their introductory

psychology course.

Results indicated that type of reinforcement, mode of presentation and mode of responding did not significantly affect achievement scores. Significantly lower error rates on the programed lessons were found for those subjects who used the programed texts, as well as for those who were provided with the correct answer, but their achievement scores were only slightly higher.

The authors hypothesized that the absence of any significant effect attributable to providing knowledge of results may have been due to the ability of students to obtain feedback from the succeeding frames being studied, or that perhaps the group which did not receive feedback "may have been forced to respond more actively to the program" thereby counterbalancing any positive effects of feedback (p. 417).

Rosenstock, Moore, and Smith (1965) continued to investigate the possible effects of feedback, this time using programed texts on set-theory, in which all the review or practice frames were eliminated. Results nevertheless again showed no significant achievement differences between 100%, partial, and no feedback of the correct answers. And again, as found by Krumboltz and Weisman (1962), this was despite the significantly

lower error rates in the group which received constant feedback.

Rosenstock, Moore, and Smith (1965) interpret their results as supportive of the arguement that feedback in the form of an answer frame in programmed instruction does not have the reinforcing properties attributed to it, and hypothesize that knowledge of results may be reinforcing "only if the learners motivation is intrinsic to the task to be learned" (p. 540). Reflecting upon the lower error rates when feedback was available within the text, they acknowledged that this "undoubtedly reflects copying of the answers by some of the Ss", but they concluded that "if copying does occur, thereby decreasing the error rate, achievement does not necessarily diminish" (p. 540).

Gilman (1969)conducted a study comparing several feedback methods for correcting errors in an attempt to differentiate between the reinforcing and informational aspects of feedback. The program utilized was designed to teach concepts in general science that are often misunderstood, thereby creating a higher than usual error rate, and providing what was believed to be a better format for investigating the effectiveness of feedback in correcting learner errors.

Seventy-five college students were assigned randomly to one of five experimental treatment groups. The program required thirty multiple choice selections and was presented via computer-assisted instruction. This study differed from most others in that the instructional program was repeated as necessary, with all incorrect frames readministered until 100% criterion was attained.

Results indicated no significant achievement differences on an immediately administered post-test between group (A) receiving no feedback, group (B) receiving feedback of 'correct' or 'wrong', group (C) receiving feedback of letter denoting the correct response choice, and group (D) which received feedback in sentence form repeating the correct answer. Only group (E) which received a combination of feedback modes, including confirmation, knowledge of the correct response, and the repetition of the full answer in sentence form showed a significant gain in achievement. An analysis of the reported group mean scores indicated that this difference was very slight, accounting for only a minimal amount of variance.

Results did indicate that learning rate, based on number of correct responses and number of iteration of the

program to criterion, was significantly improved when some form of information as to the correct answer was provided as opposed to when no feedback or simple confirmation was given. Gilman (1969) concluded that this study "indicate(s) the advantages for learning which are present when the learner is provided with the correct response after making an error, and also show(s) the advantages for retention when he is provided with extensive information in feedback messages" (p.507). He maintained that the results challenged the concept of the reinforcing value of feedback, but supported the value of providing information to students in a programed instruction format. It is important to note, however, that even the advantages cited were not reflected in any achievement gains due to simple feedback, but necessitated elaborate provisions for additional information to be provided to the student.

Significant Feedback Effects

The earliest studies in the effects of feedback were cited by Lumsdaine and Glaser (1960). Three of the earliest studies by Angell, Little, and Peterson, cite significant gains in achievement when feedback was presented. However, these studies pr[illegible]d feedback on review tests rather than for instructional programs,

and are therefore not analogous to more recent research.

Meyer's (1960) report was the only early study which utilized manipulation of feedback within an instructional, programed textbook. It is cited by Anderson, Kulhavy, and Andre (1971) as the only experiment, other than their own, which to their knowledge "has ever found that KCR facilitates learning from a programmed lesson" (p. 153).

Meyer (1960) provided each of 58 students in eighth grade classes with a programed workbook teaching common prefixes. Feedback consisted of providing the correct answer on the following page. The author noted that "close supervision reduced the tendency to peek" (p. 232), and that data was deleated when cheating was evident. Comparison groups included a "no Feedback" group, who were given feedback by the teacher on the following day; a feedback group where students placed an "X" on the wrong answers as they proceeded through the lesson; and a second feedback group which, after completing the lesson, went back through their books as many times as necessary, until all items were correct.

Results indicate a significant effect "of immediate self-scoring versus delayed 'teacher' scoring"(p. 233). These significant comparisons were of post-test gain scores

of the "no feedback" group to both "feedback" groups combined. Since no significant differences were found between the two feedback groups no other comparisons were reported. (Statistical procedures are not clearly outlined, but appear to be multiple t tests). It is also noted that the "no feedback" group made significantly more program errors than the self-scoring "feedback" groups.

More recently, Anderson, Kulhavy, and Andre (1971) report significant achievement benefits accrued from feedback. The author hypothesized that previous research failed to find significant achievement gains due to knowledge of correct results, because the programed texts used often allowed for "cheating" by the students,that is they made it possible for students to copy the correct answer. They cite a study by Anderson (1969) wherein more than 40% of the students said they had sometimes copied the answers which had been provided on the following page, and where achievement was a decreasing function of peeking. They reasoned that this may have resulted in a short circuiting of attention when the correct answer was available, in that the student may have copied the answer without reading the material in the frame, and without studying it. The 1971 study was designed to investigate this hypothesis, as well as to investigate whether knowlege-of-results

functions as a reinforcement or as corrective feedback. No significant results were reported as to how knowledge-of-results works, but a significant effect for reinforcement was found.

Anderson, Kulvavy and Andre (1971) presented an adaptation of Tobias (1968) modification of a linear program on the diagnosis of myocardial infarction from electrocardiograms (originally prepared byFrancis Mechner), on a computer based instructional system. Subjects were college students enrolled in an educational psychology course, who were stratified into three levels of verbal ability and, within each level, randomly assigned to experimental groups . The program was completed within a two hour session and required subjects to respond to multiple-choice and completions questions, both on the program and criterion post-test .

Approximately twenty subjects were assigned to each of the eight feedback conditions. Of primary interest were the significant differences in achievement scores obtained between the group receiving knowledge-of-correct -response after every frame and the group not receiving feedback. A replication of this study again showed significant achievement gains due to feedback.

In an attempt to investigate their original hypothesis that copying of answers, when possible, is deterimental to achievement they included within their replication research a "Peek" condition in which knowledge-of-results was presented at the same time the frame was exposed, so that students had access to the answer before responding. Results showed that this group learned significantly less than the group receiving feedback, and even slighty less than the group receiving no feedback (this comparison was not statistically significant).

Anderson, Kulhavy, and Andre (1971) state that their most important finding was that "students who received KCR after every frame performed better on the criterion test than students who received no KCR" (p. 153). They present two hypotheses to explain the lack of positive benefits for feedback reported in most of the previous research. Their major contention was in many of the previous studies it was possible for the student to have copied the correct answer into the blank without studying the frame. Their finding that the "Peek" group did learn substantially less than the group receiving feedback, is interpreted as supportive of this hypothesis. They explain the lack of positive results in other studies where computer terminals were used (Feldhusen & Birt, 1962; Hough & REvsin, 1963;

Moore & Smith, 1961,1964) this way:

> Our best guess is that KCR failed to show to advantage because the programs used in these studies contained many copying frames and were otherwise heavily prompted. (p.154)

Anderson, Kulhavy, and Andre (1972) designed a follow-up study to test the preceeding hypothesis. Utilizing a different program, one dealing with population genetics (Faust, Anderson, Guthrie, and Dranz, 1969) they varied both the structure of the program: copying frames and standard frames, and the presentation of the correct answer: 100% and 0%.

Results were equivocal. There was no statistically significant difference on post-test scores between the group receiving the copying version, where almost every frame was turned into a copying frame, and the standard version ($.10 > p > .05$); the 100% standard group did perform better than the 100% copying group. Additionally, there was no replication of their previously reported significant learning gains due to providing immediate knowledge of correct results; in this study the difference was only marginally significant ($.10 > p > 05$).

Summary

The bulk of evidence from research seems to indicate that feedback has no, or only slight effectiveness, in

improving learning from programed instruction.

Anderson, Kulhavy, and Andre (1971) concluded that one of the key factors involved providing feedback in such a manner as to preclude cheating and the possible short circuiting of attention. Although this appears to be a possible factor, it cannot be the only variable involved since previous research conducted on computer terminals, or machines, where cheating was clearly impossible, also reported no significant gains in learning when feedback was provided (Feldhusen & Birt, 1962; Hough & Revsin, 1963 ; Moore & Smith, 1961, 1964). Eight studies which separately compared the presentation of instructional programs via machines as opposed to textbook format, and not one found any significant differences in achievement gains (Goldstein, Gotkin (1963). Evidently, eliminating the copying of the correct answers by the use of machines or computers, does not alter final achievement scores.

Anderson, Kulhavy and Andre (1971), were aware of the inability of their cheating hypothesis to cover all contingencies, and further guessed that over-prompting within the frames of the programs used in previous research may have also caused short circuiting of attention, i.e., the students may have felt it unnecessary to read

the frames completely, thereby negating the positive effects of feedback. However, their (1972) follow-up study failed to confirm this hypothesis, and furthermore, did not replicate their previous results. All results were in the predicted direction, but were of only marginal significance.

AN ATI HYPOTHESIS

The interaction of individual differences in aptitudes and instructional treatments was postulated by Lee Cronbach in his 1957 presidential address to the American Psychological Association. At that time he strongly advocated that instructional treatments be designed to fit, not the average person, but groups of students with specific aptitude patterns. Research aimed at delineating particular aptitude treatment interactions has since proliferated.

Tobias(1973a,1976) hypothesized an interaction between the students prior achievement and the optimum level of instructional support. He suggested that the lower the level of prior achievement, the more instructional support required, and conversely, that the higher the level of prior achievement, the less instructional support necessary.

The level of prior achievement, or pre-familiarity, is defined by pretest scores, which are simple to measure and which can be task and time specific without incurring

conceptual or administrative difficulties. Instructional support implies the utilization of various techniques which are deemed to be an aid to the student, such as organization of the instructional content, presentation of instructional objectives, and providing opportunities for overt responding. Feedback, in the form of providing the student with the correct response, is also considered a form of instructional support. The implication is that the need for feedback support is inversely related to the subject's prior achievement level, i.e., feedback may benefit only the achievement of students with limited prior experience with the content.

Research on Achievement Treatment Interactions

Research studies have generally upheld the hypothesized interaction of prior achievement and optimum level of instructional support. Types of instructional support which have been experimentally manipulated and studied include sequencing (Tobias, 1971), audio tape rewind (Deutsch &Tobias, 1990; Tobias & Redfield, 1980), and constructed response (Abramson & Kagan, 1975; Tobias, 1968; Tobias & Abramson, 1971; Tobias & Ingber, 1976; Tobias & Litwak, 1977). Feedback, as a main effect variable, was not the primary focus of attention in any of these

studies, but Tobias and Abramson (1971) did include an experimental manipulation of the availability of feedback when combined with constructed responding.

The Tobias and Abramson (1971) study was designed in part to investigate the interaction of response mode and subject matter familiarity. Subjects were randomly assigned to one of three presentation groups: a reading group and two constructed response groups, one which received feedback of the correct answer, and one for whom feedback was eliminated entirely. Materials were designed so that confirmation for one frame could appear in the left-hand margin of the succeeding frame on the next page.

The program utilized was developed by Tobias (1968) and contained both familiar and technical sections. The familiar portion of the program dealt with the incidence of heart disease, and risk factors for contracting heart disease; the technical portion dealt mainly with the diagnosis of myocardial infarction from the electrocardiogram, and contained technical terminology.

The authors concluded that there was a significant interaction between response mode and familiarity. Results were that the group which received the most instructional support, i.e., both constructing responses and receiving

of the correct answers, did significantly better than the reading group, but only on the technical material. They also reported that even when constructed responding was held constant,the removal of reinforcement (i.e., feedback) lead to lower mean scores on the technical data. There was no significant differences between the three response mode groups on the familiar material.

A possible interpretation of the finding is that both constructed responding and feedback each serve as instructional support, and lead to increased achievement only on technical, non-familiar material where prior achievement levels are low. Feedback may be less important when the material to be learned was more familiar to the student.

Abramson and Kagan (1975) replicated the investigation of the interaction between prior familiarity and response mode on programed materials by experimentally manipulating familiarity. They also used the materials developed by Tobias (1968) with a reading version or a programed instruction format which included both constructed response and feedback. There was no attempt to study feedback separately from constructed responses.

Programs were administered to sixty college students, half of whom were pre-familiarized with the material to raise prior achievement levels. They

found that constructed response with feedback led to superior achievement compared to achievement from a reading mode when the material was unfamiliar, replicating previous findings. Additionally, a disordinal interaction was found, such that familiarization lead to lower achievement with constructed responding and to higher achievement in the reading condition. The authors suggested that the longer period of time taken to complete the program combined with the forced responding, may have made subjects who were already familiar with the material less attentive to the program.

Tobias and Ingber (1976) conducted an additional test of the familiarization hypothesis using a different program and younger subjects than previously employed. A linear program was devised studying the rituals involved when the five books of Moses are read during the prayer services in synagogue and some of the Hebrew words applied to the rituals. This subject matter was selected because of the great variability in familiarity among students which was independent of general intellectual level.

Subjects were 104 students recruited from a Catholic and Jewish parochial school. Results confirmed the major hypothesis: instructional support in the combined form of

constructed responding and feedback yielded superior achievement at low pretest levels, with differences becoming smaller as pretest scores increased. These results were basically replicated by Tobias and Litwak (1977).

These studies (Abramson & Kagan, 1975; Tobias & Abramson, 1971; Tobias & Ingber, 1976; and Tobias &Litwak, 1977) all found that constructed responding in conjunction with feedback interacted wtih familiarity. The extent to which either constructed response or feedback could account for the results, could not be determined.

Rationale

A re-analysis of previous research indicates the viability of Tobias(1973a,1976) hypothesis as an explanation for the lack of achievement gains that have been reported when feedback was made available. With the exception of the Anderson (1971) study, they all appear to have dealt with subject matter which was at least somewhat familiar to their subjects.

Moore and Smith (1961) used a program teaching spelling to sixth graders, a subject matter which can be presumed to be relatively familiar to these students. Krumboltz and Weisman's (1962) study involved teaching prospective teachers how to interpret educational test results, an area with which prior familiarity probably

varied among subjects. Hough and Revsin (1963) utilized a program teaching college students in the department of secondary education the historic foundations of the secondary school. These programs involved subject matter of some familiarity to their respective subjects.

Most interestingly, Feldhusen and Birt (1962) specifically selected their topic (teaching machines and programed instruction) because they believed it to be an area free from previous learning for their subjects (general psychology students). Pretest scores are available in this instance, and the average score was four out of fifteen correct on a constructed response format. This seems to indicate that even here the material was not novel to most of the students.

In the more recent studies Olson (1972) taught the importance of color to undergraduate interior design students, and Oner (1972) used the Sullivan Programmed Math Series for sixth graders. Both used teaching material which intuitively appears to be of a nature familiar to the subjects.

Goldstein and Gotkin (1961) reviewed eight different studies which compared the mastery of subject matter when the identical program was presented by machine or text-book format, and found no significant differences due to mode of presentation. As noted previously, this essentially

indicated no differences in achievement when feedback follows rather than preceeds or co-occurs with the response. An attempt to reinterpret these results in terms of the Tobias(1973a,1976) hypothesis is difficult since access to the original studies was not possible. But the information available seems to indicate little diversity in terms of students' prior familiarity with content.

Two of the projects utilized programs teaching spelling to fourth graders, two involved teaching binary numbers to college students, another three taught various mathematical principles to high school and college students, and the eighth taught basic electricity to telephone technicians. It appears that all programs were similar in that the general area of the material presented, and presumably some of the specific content, was one with which the subjects had some prior familiarity.

Whereas all the above studies cited found that manipulating feedback did not significantly increase achievement results, none of them utilized highly novel subject matter in their respective programs of instruction. There is indication, however, that when completely novel material is the subject matter involved, feedback does lead to a significant increase in learning.

When Anderson, Kulhavy and Andre (1971) used a program on the diagnosis of myocardial infarction, an

area where prior achievement scores were almost nil, feedback was found to lead to a significant increase in achievement. It is of significance that the same team (Anderson, Kulhavy & Andre, 1972) did not obtain statistically significant results when the program was changed. They report that "The standard program consisted of the first 104 frames adapted from a programmed introduction to population genetics (Faust, Anderson, Guthrie & Dranz, 1969). The subject matter was unfamiliar to most of the potential subjects, and the program is known to teach effectively." (p. 187) No other specifics about the program are given. It remains plausible that this program dealth with material more familiar to their college students than the program on myocardial infarction used in the 1971 study, and that this accounts for the fact that results were only statistically of marginal significance.

It is contended that the critical variable accounting for the increased effectiveness of feedback in the Anderson, Kulhavy, and Andre (1971) and the Tobias and Abramson (1971) experiments as opposed to no significant differences attributed to feedback procedures in other research, may be the differential degree of prior familiarity and prior achievement levels with the specific material to be learned. This study was therefore undertaken to

specifically investigate the hypothesis that feedback in programmed instruction in an important variable in the learning of novel, but not familiar subject matter.

METHOD

The research design consisted of the experimental manipulation of the feedback variable, with the assigned variable, level of prior achievement, determined on the basis of pretest score; post-test scores were used as the dependent achievement measure.

Materials

The program utilized was an updated version of one that was developed by Tobias & Ingber (1976). It deals with the rituals involved when the five books of Moses are read during the prayer services in synagogue and includes some of the Hebrew words applied to these rituals. This subject matter was selected because of the wide variability in familiarity among students about these religious rituals which is not directly related to other intellectual or scholastic achievements.

The program contained 41 frames, and required 72 constructed responses. All subjects received the linear program in booklet form, one frame per page. Each booklet contained detailed instructions as well as a sample question,

in order to assure that directions were understood. The answer sheets provided, all contained space for recording approximately 18 answers per column.

Feedback availability was experimentally varied in three ways. The "No Feedback" group was given the program in the standard manner, with no answers provided in either the booklets or answer sheets. The "Peek" group was provided with answers which were placed on the lower right hand corner of each frame of the program, so that the feedback could be viewed before the answer was given. The "Feedback" group was provided with the correct answer, which was made available only after they had responded, by the use of specially designed answer sheets. Each answer sheet contained two sheets of NCR paper attached together at both sides, and precut between questions. Subjects were instructed to record their answer on the top sheet which was blank, and then to tear off and discard the top tab on which they had written. This revealed the bottom sheet containing the correct answer as well as a carbon copy of their hand written answer.

The pretest and post-tests were identical and were administered to all groups. It contained 44 blanks which were to be filled in on the question sheet. The alpha reliability of this test was reported to be .94 (Tobias & Ingber, 1976).

Subjects

Subjects were boys and girls ages 10 to 11. Eighty four subjects were recruited from a Jewish day camp and forty eight from a Hebrew afternoon school and randomly assigned to treatments. Both settings were selected because the children attending varied greatly in their prior familiarity with the program material.

Procedure

The pretest was administered to subjects in their natural setting in the week preceeding the actual experimentation. They were simply told that this was part of a survey and were asked to cooperate.

On the following week, with interval time ranging from five to seven days, all subjects were randomly assigned to one of three groups and separated into different rooms or areas.

The experimental treatment groups were led by their leaders, counsellors and teachers, who were randomly assigned to rooms. They received specific oral and written instructions, and were closely supervised by the experimenter.

The following general instructions were given by their respective leaders to all groups:

> This lesson is presented in a form called 'Programmed Instruction', and deals with one area of Jewish customs. The lesson is presented in a series of boxes, or frames, each of which require one or more answers.

Each group was then given the instructions specific to their treatment group, orally by their leader and in written form in their respective booklets. Program answer sheets and booklets were collected as each subject finished. Time to completion was recorded. Post-tests were then administered. Game sheets were provided to students when they handed in the post-tests, in order to fill-in differences in individual work speeds.

At the completion of the experiment the nature of the research was explained to both subjects and leaders.

RESULTS

Data was analyzed by multiple linear regression using a stepdown procedure involving an initial ordering of effect, and estimation of each effect adjusted for those preceding it in the ordering and ignoring those so following it. (Overall & Spiegel, 1969). The dependent variable was the post-test achievement scores. The independent variables were tested in the following order: pretest scores, feedback treatment group, sex of subject, site of testing (camp or school).

The basic regression formula for main effects was $y = B_o + B_1Pre + B_2Tr_1 + B_3Tr_2 + B_4Sex + B_5Site + e$

with: y = post-test scores as a continuous variable

B_o = constant

Pre = pretest scores as a continuous variable

Tr_1, Tr_2 = experimental treatment group; coded "0,1" "1,0" and "0,0" for feedback, no feedback, and peek groups respectively

Sex = sex of subject; coded "1" for males, and "0" for females

Site = site of testing; coded "1" for school subjects, and "0" for camp subjects

e = unexplained variance

The results summarized in Table I indicated that pretest effects accounted for 46% of the variance. Treatment involving the experimental manipulation of feedback, was also highly significant ($p < .01$) and accountd for 4% of the variance. The "No Feedback" group scored slightly higher than the "Feedback" group, and the "Peek" group scored lowest. Sex of subjects was significant ($p<.05$) with girls scoring higher than boys and accounting for 2% of the variance and unexpectedly, the school group scored higher than the camp site group ($p <.0001$) and accounted for 6% of the variance.

Interactions were sequentially added to the basic regression formula. All double interactions were tested in the presence of the main effects. The following order was used, with each interaction dropped from the equation if it was found to be nonsignificant: Pretest X Treatment; Site X Treatment; Site X Pretest; Treatment X Sex; Sex X Pretest; Sex X Site. Only "Sex X Site" was found to be significant, ($p < .05$) and accounting for 1% of the variance. The predicted interaction between "Pretest X Treatment" was not significant.

Triple interactions considered meaningful to the main hypothesis were tested in the presence of all main effects and double interactions. Those tested were "Sex X Treatment X Pretest" which was nonsignificant; and "Site X Treatment X Pretest", which was marginally significant, ($p < .06$) and accounting for 1% of the variance.

Two subsidiary analyses of the data were performed. Analysis of the School group separately showed a "Treatment X Pretest interaction accounting for 4% of the variance. But due to the small number of subjects (N = 48) the interaction was statistically nonsignificant.

Subsidiary analyses with time as the dependent measure were performed with feedback treatment, pretest scores, sex of subjects and group site as the independent measures.

Treatment effects were highly significant on time, accounting for 53% of the variance. This reflected obvious differences during testing, with the "Peek" group always finishing first, followed by the "No Feedback" group and with the "Feedback" group clearly finishing last. Pretest achievement scores had no effect on time to completion but boys finished quicker than girls ($p < .05$; 1% of the variance accounted for) and the camp sample finishing before the school site subjects ($p < .05$; 2% of the variance accounted for).

DISCUSSION

The predicted interaction between pretest scores and treatment was not significant. Both the obtained results and observational reports support the interpretation that this was due to the low motivation of the subjects in the summer camp sample.

It was evident from observation that many of the subjects in the camp sample were unhappy about being asked to do anything resembling school work, and the boys especially kept asking how much longer they had to work. The subjects in the school sample, in comparison, seemed interested in the program as a relief from their normal schedule, and boys as well as girls worked more diligently and without complaints.

Results support observational reports, in that time on task was shortest for the boys and for the camp sample, reflecting a relative lack of motivation or interest. This correlated directly with main effect differences indicating that boys performed significantly more poorly than girls, and that the camp sample had lower scores than the school sample. Moreover, pre-to post-test gains of one or less points were obtained by 23% (13 of 56 S's) of the boys at the camp site, compared to only 7% (2 of 29 S's) of the boys at the school who showed such poor gains. (The significant "Site X Sex" interaction reflects a diminishing difference between boys and girls in the school sample.)

The results of this study and the analysis of prior research suggest a more complex model regarding the effects of feedback on learning from instruction. Feedback is likely to lead to increments in learning in terms of the degree to which it enga es students to attend to the material, understand it thoroughly, and figure out relationships between the content and their prior experiences. Apparently in some cases, when motivation is high, the absence of feedback may promote much active efforts by students to thoroughly comprehend the material. The data from this study suggests that this may have been the case

in the school sample where the no feedback group obtained non-significantly higher scores.

The present analysis suggests that in prior research the non-significant findings for the effects of feedback may have been attributable to the failure of the feedback condition to stimulate intense efforts by the students to comprehend the material more fully than other modes. Anderson et al's (1971) study used the complex novel materials prepared by Tobias (1968). In several different studies (Tobias, 1968, 1973a, b) the error rates for this program approximated 25%. It seems reasonable to assume that for material of such complexity, feedback facilitated student attempts to comprehend the content thoroughly. Studies in which feedback did not facilitate learning frequently used materials with very low error rates. For example, Krumboltz and Weisman (1962) report error program error rates of 12% and Hough and Revsin (1963) report error rates of only 2%. Even in the present study the program error rates were relatively low, averaging only 15%. Analysis of these data confirm the hypothesis that feedback can be expected to improve learning to the degree to which it facilitates students' attempts to obtain a thorough understanding of the subject matter.

CONCLUSIONS AND IMPLICATIONS

This research did not fully support the interaction predicted between feedback support and prior achievement. This seemed due to the low motivation of the group, indicating that future research should be conducted within the setting for which it is meant, so that motivational factors are not a confounding factor.

This study concurs with the vast body of research which questions the benefits to be accrued from feedback. Theoretically, it implies that feedback support should be defined in terms of the extent to which it elicits students' thinking about the content.

Practically, this study serves to seriously question the benefits of providing feedback in programmed instruction material, especially when cheating is made easy, thereby curtailing the student's tendency to think about the subject matter.

REFERENCES

Abramson, T, & Kagan, E. Familiarization of content and different response modes in programmed instruction. Journal of Educational Psychology, 1975, 67, 83-88.

Anderson, R. C., The control of student mediating processes during verbal learning and instruction. Review of Educational Research, 1970, 40, 349-370.

Anderson, R. C., Kulhavy, R. W., & Andre, T. Feedback procedures in programmed instruction. Journal of Educational Psychology, 1971, 62, 148-156.

Anderson, R. C., Kulhavy, R. W., & Andre, T., Conditions under which feedback facilitates learning from programmed lessons. Journal of Educational Psychology, 1972, 63, 186-188.

Cronbach, L. J., The two disciplines of scientific psychology. American Psychologist, 1957, 12, 671-684.

Deutsch, T., & Tobias, S., Interaction among prior achievement, anxiety and instructional method. Paper delivered to the annual convention of the American Psychological Association. Montreal: September, 1980.

Feldhusen, J. F., & Birt, A., A study of nine methods of presentation of programmed learning material. Journal of Educational Research, 1962, 55, 461-466.

Gilman, D. A., Comparison of several feedback methods for correcting errors by computer-assisted instruction. Journal of Educational Psychology, 1969, 60, 503-508.

Goldstein, L. S., & Gotkin, L. G., A review of research: Teaching machines vs. programed textbooks as presentation modes. Journal of Programed Instruction, 1962, 1, 29-35.

Hough, B., & Revsin, B., Programmed instruction at the college level: A study of several factors influencing learning. Phi Delta Kappan, 1963, 44, 286-291.

Krumboltz, J. D., & Weisman, R. G. The effect of intermittent confirmation in programmed instruction. Journal of Educational Psychology, 1962, 53, 250-253.

Lumsdaine, A. A., & Glaser, R. Teaching machines and programmed learning. Washington, D. C.: National Education Association, 1960.

Meyer, S. R. Report on the initial test of a junior high school vocabulary program. In A. A. Lumsdaine & R. Glaser (Eds.), Teaching machines and programmed learning. Washington, D. C.: National Education Association, 1960, p 229-246.

Moore, J. W., & Smith, W. I. Knowledge of results in self-teaching spelling. Psychological Reports, 1961, 9. 717-726.

Moore, J. W., & Smith, W. I. Role of knowledge of results in programed instruction. Psychological Reports, 1964, 14, 407-423.

Olson, G. H. A multivariate examination of the effects of behavioral objectives, knowledge of results, and the assignment of grades on the facilitation of classroom learning. Dissertation Abstracts International, 1972, 32 , 6214-5.

Oner, N. P. Impact of teachers behavior and teaching technique on learning by anxious children. Dissertation Abstracts International, 1972, 32, 6215.

Overall, J. E., & Spiegel, D. K. Concerning least squares analysis of experimental data. Psychological Bulletin, 1969, 72, 311-322.

Rosenstock, E. H., Moore, W., & Smith, W. I. Effects of several schedules of knowledge of results on mathematical achievement. Psychological Reports, 1965, 17, 535-541.

Skinner, B. F. The science of learning and the art of teaching. Harvard Educational Review, 1954, 24, 86-97.

Skinner, B. F. The technology of teaching. New York: Appleton-Century-Crofts, 1968.

Tobias, S. The effect of creativity, response mode, and subject matter familiarity on achievement from programmed instruction. New York: MSS, Educational Publishing Co. 1968.

Tobias, S. The effect of sequence and familiarity with subject matter in achievement from programed instruction. Technical Report #6, Programed instruction research program, New York: City University, July, 1971.

Tobias, S. Review of the response mode issue. Review of Educational Research, 1973, 43, 193-204. (a)

Tobias, S. Achievement treatment interaction. Review of Educational Research, 1976, 46, 61-74. (b)

Tobias, S., & Abramson, T. Interaction among anxiety, stress, reponse mode and familiarity of subject matter on achievement from programmed instruction. Journal of Educational Psychology, 1971, 62, 352-364.

Tobias, S., & Ingber, T. Achievement treatment interactions in programmed instruction. Journal of Educational Psychology, 1976, 63, 43-47.

Tobias, S., & Litwak, D. L. Adapting instruction to achievement differences. Paper delivered to the annual convention of the American Educational Research Association, New York: April, 1977.

Tobias, S., & Redfield, R. Anxiety, prior achievement and instructional support. Paper deliver to the annual convention of the American Educational Research Association. Boston: April, 1980.

TABLE I

Summary of Multiple Linear Regression Results: Posttest Achievement is Dependent measu

Independent Measure		F value	Percent Variance Accounted For
PRETEST	***	125.511	46%
FEEDBACK TREATMENT	**	5.275	4%
SEX	*	6.013	2%
SITE	***	20.463	6%
PRETEST X TREATMENT	NS		
SITE X TREATMENT	NS		
SITE X PRETEST	NS		
TREATMENT X SEX	NS		
SEX X PRETEST	NS		
SEX X SITE	*	4.411	1%
SEX X TREATMT X PRETEST	NS		
SITE X TREATMT X PRETEST	@	2.790	1%

*** $p < .001$
** $p < .01$
* $p < .05$
@ $p < .06$

TABLE II

Means and Standard Deviations for the three Experimental Manipulations

TREATMENT	POSTTEST SCORE		PRETEST SCORE		TIME ON PROGRAM		PROGRAM ERRORS	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Feedback N=46	16.09	9.6	8.17	6.4	34.48	7.7	10.30	11.4
No Feedback N=42	18.19	12.4	6.38	5.3	24.85	7.1	21.55	22.2
Peek N=44	14.04	9.3	8.55	6.8	20.32	5.6	1.18	3.1
Total N=144	16.08	10.9	7.73	6.5	28.16	10.9	10.84	16.7

TABLE III

Means and Standard Deviations --- School Site Sample Only

TREATMENT	POSTTEST SCORE		PRETEST SCORE		TIME ON PROGRAM		PROGRAM ERRORS	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Feedback N =16	17.44	8.6	6.44	4.6	38.69	5.7	6.00	6.7
No Feedback N=15	20.60	7.0	6.60	4.2	24.86	4.2	10.13	17.0
Peek N=17	17.76	8.6	8.59	3.7	26.00	5.7	2.65	6.5
Total N=48	18.54	8.1	7.25	4.2	29.88	8.2	6.10	11.18

TABLE IV

Means and Standard Deviations ----- Camp Site Sample Only

TREATMENT	POSTTEST SCORE		PRETEST SCORE		TIME ON PROGRAM		PROGRAM ERRORS	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Feedback N=30	15.37	10.2	9.10	7.4	38.37	8.7	12.60	13.9
No Feedback N=27	16.85	15.4	6.26	5.9	24.85	8.7	27.89	25.1
Peek N=27	11.70	9.8	8.52	8.8	16.74	5.6	.25	1.0
Total N=84	14.67	12.0	8.00	7.5	27.17	12.1	13.55	19.8